Proceedings of the Iowa Academy of Science

Volume 39 | Annual Issue Article 51

1932

A Sampling Test of the Technique of Analyzing Variance in a 2xN Table with Disproportionate Frequencies

Bernice Brown

A SAMPLING TEST OF THE TECHNIQUE OF ANALYZING VARIANCE IN A 2 x N TABLE WITH DISPROPORTIONATE FREQUENCIES

BERNICE BROWN

The data used in this study are taken from the 1928 and 1929 records of the meat fed Wistar rats in the Animal Research Laboratory of the Foods and Nutrition Department at Iowa State College. The records of 149 animals were available, unequally divided as to sex, and including four generations. The weight of the animals at 28 days of age was recorded with the weight each week thereafter for six successive weeks.

For the purpose of study the rats were divided into eight groups in accordance with the sex and generation. The number of experimental animals (N) in these groups with the mean initial weight (I) and the mean gain of each group (G) is shown in Table I. Fisher's method of analysis of variance was used (1), (2), and (4).

Table I Means of Initial Weight and Gain Together With Frequencies of Eight Sex-Generation Groups

		Generation 2	Generation 3	Generation 4	Generation 5	Total
Male	I	50.6190	45.2667	37.3500	45.0000	45.6364
	G	176.9524	161.4667	155.6667	171.0000	167.3273
	N	21	15	12	7	55
Female	I	42.9630	43.9600	36.7391	43.2632	41.7660
	G	109.5185	114.0800	106.5217	108.7895	109.9362
	N	27	25	23	19	94
Total	I	46.3125	44.4500	37.0857	43.7308	43.1946
	G	139.0208	131.8500	124.6857	124.0769	131.1208
	N	48	40	35	26	149

I, Initial Weight; G, Gain; N, Number.

It is known that the disproportion of the frequencies between the sexes in the several generations distorts some of the estimates of the variance as obtained from the sample. Hence the method of adjustment developed by Brandt (3) was used.

Samples were drawn in which the sex generation groups were equally represented. This was done in the hope that in creating a properly designed experiment we would be able to set up empirical values of variance between sex, variance between generations, and variance due to interaction, with which we could compare the

estimates of variance obtained by making the adjustment for disproportionate frequencies. The samples were drawn at random without replacement, the only restriction being that equal numbers should be drawn from each sex generation group. Forty successive samples were drawn.

Two limitations of the sampling are noted here. The drawings were made without replacement because from each sample lots were to be selected. The methods of selecting lots commonly used by research workers in nutrition are necessarily based on this type of sampling. A further limitation of the sampling method arises from the fact that there is one sex generation group (5 male) which contains only seven observations. In drawing samples of 56 observations each it is apparent that every sample contained this entire group. It happens also that it is a widely varying group, with a range in initial weight from 39 to 60 grams and a range of weight gains from 154 to 203 grams. The probable effect of this is a slight increase of the variance within the groups with an accompanying lowering of the apparent significance of differences between groups.

Table II exhibits the results. The first column shows the estimates of variance as obtained from the 40 samples. These are average estimates per sample derived from the analysis of variance. Although the sampling was not extensive and the method of sampling had the limitations referred to, we regard these estimates as the best estimates we have of the variance in a population in which the sex generation groups are equally represented.

Table II. Estimates of Variance

Sources of Variation	Estimates of Variance in Population		
		Adjusted Values	
	40 Samples	Interaction Only	Brandt's Formulae
a. Initial Weight			
Total	113.6185	117.4564	117.4564
Within	95.5758	105.0224	105.0224
Between	128.8107	138.2766	138.2766
Between sex	189.5351	195.4905	159.4309
Between generation	230.0337	230.9543	216.5016
Interaction	57.3463	38.8304	46.7356
b. Gain			
Total	1241.5981	1194.8373	1194.8373
Within	396.7345	409.1824	409.1824
Between	6765.6959	6396.8420	6396.8420
Between sex	44993.3297	42953.4168	43375.5128
Between generation	407.2345	721.1264	208.5014
Interaction	381.3770	398.7258	706.3896

The second column shows the results obtained by applying Brandt's adjustment to the interaction estimate only. In this case the adjustment is used only for the purpose of eliminating the spurious interaction due to the disproportionate frequencies leaving us an estimate of the true, biological interaction. These are all fair approximations of the variability in a population in which the groups are represented by equal frequencies. In the case of initial weight, the estimate of variance due to interaction is the poorest approximation but with the gain factor the variance between generations is the poorest.

The third column in Table II shows the estimates derived by using all of Brandt's formulas. In the case of initial weight, this method of making the analysis of variance has given us a better approximation of the estimate of variance due to interaction. It is remarkable that from the rather serious handicap caused by the disproportionate frequencies in the original sample, we are able by the algebraic adjustment to get such a good estimate of the true interaction.

With regard to the gain factor, the estimate of variance between sex is good and the estimate of variance between generations is a better approximation than we were able to get in column two without the adjustment. However, the estimate of variance due to interaction is a poor approximation.

It is perhaps sufficient to say that none of the differences in any set of estimates are large enough for significance.

No conclusions can be drawn as to the efficacy of the adjustment or the relative merits of the method of making the analysis of variance. Plans are being made to carry on more adequate sampling in the hope of getting positive results.

REFERENCES CITED

1. FISHER, R. A. Statistical Methods for Research Workers. Fourth Edition. Oliver and Boyd, Edinburgh. 1932.
2. FISHER, R. A., and MACKENZIE, W. A. The Manurial Response of different Potato Varieties. Journal of Agricultural Science, 13:311. 1923.
3. BRANDT, A. E. Analysis of Variance in a "2xs" table with disproportionate frequencies. Journal of the American Statistical Association, 28:164. 1933.
4. SNEDECOR, GEORGE W. Computation and Interpretation of Analysis of Variance and Co-variance. Iowa State College Official Publication. 1933.